

1890

1891

1892

1893

1894

1895

1896

1897

1898

1899

1900

1901

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 30, 2002, 19:51:03 ; Search time 2499.58 Seconds
(without alignments)
11483.966 Million cell updates/sec

Title: US-09-668-119-1

Sequence: 1 atgaggaaagctgtgtgac.....cacatcttcacgacatga 1740

Scoring table: OLIGO_NUC
Gapex 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb.ba:*
2: gb.hg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.om:*
20: em.or:*
21: em.ov:*
22: em.pat:*
23: em.ph:*
24: em.pl:*
25: em.ro:*
26: em.sts:*
27: em.sy:*
28: em.un:*
29: em.vi:*
30: em.hggo.hum:*
31: em.hggo.inv:*
32: em.hggo.rod:*
33: em.hggo.hum:*
34: em.hggo.inv:*
35: em.hggo.rod:*
36: em.hggo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1740	100.0	3334	9	AF056191	AF056191 Homo sapi
2	1092	62.8	3170	9	AF328769	AF328769 Homo sapi
3	962	55.3	1048	9	HS080745	HS080745 Homo sapi
4	549	31.6	2106	9	BC007529	BC007529 Homo sapi
5	304	17.5	64231	2	AL359452	AL359452 Homo sapi
6	304	17.5	157080	2	AL358856	AL358856 Homo sapi
7	304	17.5	163908	9	AC007050	AC007050 Homo sapi
8	250	14.4	145356	9	AC004033	AC004033 Homo sapi
9	203	11.7	4701	9	AK000003	AK000003 Homo sapi
10	137	7.9	64231	2	AL359452	AL359452 Homo sapi
11	132	7.6	386	6	AX071270	AX071270 Sequence
12	11	4.9	236	11	AF021108	AF021108 Homo sapi
13	86	4.5	192592	9	AC005500	AC005500 Homo sapi
14	79	4.8	107642	10	AF125313	AF125313 Mus muscu
15	45	2.6	164	11	PCU64652	PCU64652 Paracharter
16	45	2.6	307	11	G64292	G64292 6 Human Hom
17	45	2.6	383	9	HS091A	HS091A H. sapiens O
18	45	2.6	632	9	AF223572	AF223572 Homo sapi
19	45	2.6	920	9	HSTFIIDA	HSTFIIDA H. sapiens L
20	45	2.6	942	9	PPA133270	PPA133270 Pan panis
21	45	2.6	1020	9	HUMTFIIDA	HUMTFIIDA Homo sapien
22	45	2.6	1310	6	AX002276	AX002276 Sequence
23	45	2.6	1310	6	E28256	E28256 Purificatio
24	45	2.6	1475	9	HUMRNAF	HUMRNAF Human (clon
25	45	2.6	1710	9	HSTFIIDX	HSTFIIDX H. sapiens.T
26	45	2.6	1876	9	HUMTFIID	HUMTFIID Human TATA
27	45	2.6	3263	6	AX002278	AX002278 Sequence
28	45	2.6	3263	6	E28258	E28258 Purificatio
29	45	2.6	4286	6	AX002277	AX002277 Sequence
30	45	2.6	4286	6	E28257	E28257 Purificatio
31	45	2.6	113820	9	HS191N21	HS191N21 Human DNA
32	45	2.6	157115	9	AC006512	AC006512 Homo sapi
33	45	2.6	171288	2	AC019343	AC019343 Homo sapi
34	45	2.6	179289	2	AC024883	AC024883 Mus muscu
35	45	2.6	222930	9	HSU47924	HSU47924 Human chrom
36	44	2.5	1654	10	MMPOU2	MMPOU2 M. musculus
37	44	2.5	3200	10	RATARB	RATARB Rat androge
38	44	2.5	3217	6	IO9480	IO9480 Sequence 2
39	44	2.5	3864	10	MUSPOUDOMB	MUSPOUDOMB Mouse brain
40	44	2.5	4135	6	IO9511	IO9511 Sequence 3
41	44	2.5	4137	10	RATANDREC	RATANDREC Rat androge
42	44	2.5	202807	10	AC021063	AC021063 Mus Muscu
43	43	2.5	3166	3	DROENGA	DROENGA D. virilis e
44	43	2.5	217616	2	AC079531	AC079531 Mus muscu
45	42	2.4	163	4	BOVTBPR	BOVTBPR Bos taurus

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS
1	AF056191	Homo sapiens tPA inducible protein mRNA, complete cds.	3334 bp	1	AF056191.1	GI:3037134	human.	11-OCT-2000	Abraham, S. and Solomon, W.B.	A novel glutamine-rich putative transcriptional adaptor protein (TIG-1), preferentially expressed in placental and bone-marrow tissues	Gene 255 (2), 389-400 (2000)	11024300	2 (bases 1 to 3334)	Abraham, S. and Solomon, W.B.
							human.							
							Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
							Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.							
							Abraham, S. and Solomon, W.B.							
							A novel glutamine-rich putative transcriptional adaptor protein							
							(TIG-1), preferentially expressed in placental and bone-marrow							
							tissues							
							Gene 255 (2), 389-400 (2000)							
							11024300							
							2 (bases 1 to 3334)							
							Abraham, S. and Solomon, W.B.							


```

Db      689  CGTCCCGCAGCCGCGCAGCCGCTCACAGCCCACTCCAAAGTCAGCTCGGCCCTG 748
QY      1223  CCCCATCCCCAGTAGTCTCTGCCCCCAGCCCTCAGCCAGCCCTCCAGAGCCCAATGA 1282
Db      749  CCCCATCTCCAGTAGTCTCTGCCCCCAGCCCTCAGCCAGCCCTCCAGAGCCCAATGA 808
QY      1283  CAGCGCGAGCCCAAGAACCTAGTGTCCCTCAGCTGACCTTAAACACACCTGTGA 1342
Db      809  CCGCGCGAGCCCAAGAACCTAGTGTCCCTCAGCTGACCTTAAACACACCTGTGA 868
QY      1343  acccagctctgcatgagccagctgagctccagccagctgaggaagcagctgagc 1402
Db      869  ACCCAGCTCTGTCATGAGCCCGAGCCGCTCCAGCCAGCTGAGAGCAGCAGTACCTGG 928
QY      1403  acaagctgagcagcagctgctgagatcatgagccctgagccgcatgatacaagatcg 1462
Db      929  ACAAGCTGAAGCAGCTCTGCAAGTACATCGAGCCCGCTGCCCGCATGATCAACAAGATCG 988
QY      1463  acaagaacagagacagaaaaagacacctgagtaagatgaaagacaccttctgagac 1515
Db      989  ACAAGAACAAGACAGAAAAAAGACCTGAGTAAGATGAAGAGCCTTCTGTGAC 1041

```

```

RESULT  4
LOCUS   BC007529      2106 bp      mRNA      PRI      12-JUL-2001
DEFINITION BC007529 Homo sapiens, clone IMAGE:3350171, mRNA, partial cds.
ACCESSION BC007529
VERSION   BC007529.1 GI:14043090
KEYWORDS
SOURCE   human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 2106)
AUTHORS   Strausberg, R.
TITLE      Direct Submission
JOURNAL    Submitted (10-MAY-2001) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA

```

```

REMARK    NIH-MGC Project URL: http://mgc.ncl.nih.gov
          Contact: MGC help desk
          Email: cgaabs-remail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Rubin Laboratory
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Genome Sequence Centre,
          BC Cancer Agency, Vancouver, BC, Canada
          info@bcsc.bc.ca
          Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
          Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
          Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
          Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
          Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
          Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Scott,
          Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
          George Yang, Scott Zuyderduyn, Marco Marra.

```

```

FEATURES
source
Location/Qualifiers
1..2106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3350171"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH-MGC_16"
/lab_host="DH10B-R"

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov> Series: IRAL Plate: 22 Row: n Column: 21 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7209306.

```

CDS
</note="Vector: potB7"
<1..1157
/codon_start=-3
/product="Unknown (protein for IMAGE:3350171)"
/protein_id="AA07529.1"
/db_xref="GI:14043091"
/translation="TAAVAIPSSSIPILGRPMVAQVSSSLPMLSSPSPGOVONPQSM
PPPPPPPPPPPPSSSSPPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
GPLNTPVNPSVSSVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
KMSLSDITLDPKRCPLKTLQCEIALEKLNKMAVPPPPPPPPPPPPPPPPPPPPPP
DAVLANTRSPEVFNHSLYRFEVPMATLHGPPIAIPVCTRRRLLEDERSISVLOG
EVARLDPKFLVNDPSHSCSNNGVTHLICKDDKDLPSVPLETSVPADYPAOSPLMTD
ROMOYDANPFLQSVHRCMSTRLLQLPKHSVTALNLTWAGSVHQAICLSAA"

```

```

BASE COUNT 456 a 719 c 553 g 378 t
ORIGIN

```

```

Query Match 31.6%; Score 549; DB 9; Length 2106;
Best Local Similarity 99.9%; Pred. No. 3e-259;
Matches 669; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

QY      1072  caggtcagcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1131
Db      60  CAGGTCAAGCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 119
QY      1132  acccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1191
Db      120  ACCCCGCAAGTGAAGCCCCCTCCCGCCGCTCCCGCAGCCGCGGCGCAGCCAGCTCA 179
QY      1192  cagcccaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1251
Db      180  CAGCCCAACTCCAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 239
QY      1252  ccctcaccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1311
Db      240  CCTCAGCCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTCTC 299
QY      1312  cccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1371
Db      300  CCTCAGCCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTCTC 359
QY      1372  tccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1431
Db      360  TCCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTCTC 419
QY      1432  gagccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1491
Db      420  GAGCCCTCTGCGCCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 479
QY      1492  agtaagatgaaagacacctctgagatctctgagacacctctgagacacctctgagac 1551
Db      480  AGTAAGATGAAGACCTCTTGTGACATTTGTGACACACCTCTGAAAGCGGTGCTGGAAG 539
QY      1552  acccttgaaaaagtgatgagcagcagcagcagcagcagcagcagcagcagcagcagc 1610
Db      540  ACCTTGCAAAAAGTGTGAGATGCGCTTGAGAACTCAAGATGACATGAGCGGTGCCACT 599
QY      1611  ccccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1670
Db      600  CCCCCAGCCGCGCGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 659
QY      1671  gccctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1730
Db      660  GCGCTCTGCGCAATCCGCTCACTGCTTCAACCATTCCTCTGACCGCAGCATTCGTT 719
QY      1731  ccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1740
Db      720  CCAGCCATGA 729

```

```

RESULT  5
LOCUS   AL359452      64231 bp      DNA      HTG      13-JUN-2001

```



```

|||||
Db 75386 TTCAGAGGACCGTGTGAGCAGGACACACCTCCGCTCCAGGCTGTGCCAGCAGC 75327
Qy 791 tgcagcagatgcatcacaacagcagccacagccacagccacagccacagccacag 850
Db 75326 TGCAGCAGATGTCATCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 75267
Qy 851 ttgcctagaagcaacacatcacacacccagccagccagccagccagccagccagccag 910
Db 75266 TTGCTCAGAACCAACCATCACAACCTCCGCGCAGCTGCGAGACCCAGCTTGTGTGTCAC 75207
Qy 911 aggcgcagacgtctctctgacacaaatgtttatatacccaaccacacactgaattgt 965
Db 75206 AGGCGCAAGCTCTCCCTGAGCAAAATGTTGTATATACCAACCACTGAATTTGT 75152

RESULT 7
AC007050 163908 bp DNA PRI 31-MAY-2000
DEFINITION Homo sapiens chromosome 22q11 clone bac32, complete sequence.
ACCESSION AC007050
VERSION AC007050.25 GI:6456174
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Homo sapiens Chromosome 22q11 BAC Clone bac32 In BCR12-GGT Region
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Emanuel, B., Budarf, M. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 163908)
REFERENCE 3 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 163908)
REFERENCE 4 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (07-JUL-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
5 (bases 1 to 163908)
REFERENCE 5 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (22-JUL-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
6 (bases 1 to 163908)
REFERENCE 6 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
7 (bases 1 to 163908)
REFERENCE 7 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
8 (bases 1 to 163908)
REFERENCE 8 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

```

```

REFERENCE 9 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
10 (bases 1 to 163908)
REFERENCE 10 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-1999) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
11 (bases 1 to 163908)
REFERENCE 11 (bases 1 to 163908)
AUTHORS Chen, F., Hu, X., Budarf, M., Emanuel, B. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (31-MAY-2000) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Nov 20, 1999 this sequence version replaced gi:6067157.
Because these overlapping clones came from different libraries
there are numerous instances of insertions, deletions, and single
nucleotide polymorphisms in the overlapping regions below.
AC004033 (p.m11) 89227 145356 (0) overlaps AC007050(bac32) 1 59304
AC004033 (p.m11) 89227 145356 (0) overlaps AC007050(bac32) 1 59304
AC007308 (pac408) 1 22675 (193113).
FEATURES
    source
        1..163908
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="22q11"
            /clone="bac32"
BASE COUNT 39327 a 41642 c 42088 g 40851 t
ORIGIN
Query Match 17.5% Score 304; DB 9; Length 163908;
Best Local Similarity 99.7%; Pred. No. 9.5e-139;
Matches 354; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 611 agatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 670
Db 14559 AGATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 14618
Qy 671 aacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 730
Db 14619 AACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 14678
Qy 731 ttcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 790
Db 14679 TTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 14738
Qy 791 ttgcagcagatgcatcacaacagcagccacagccagccagccagccagccagccagccag 850
Db 14739 TGCAGCAGATGTCATCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 14798
Qy 851 ttgcctagaagcaacacatcacacacccagccagccagccagccagccagccagccag 910
Db 14799 TTGCTCAGAACCAACCATCACAACCTCCGCGCAGCTGCGAGACCCAGCTTGTGTGTCAC 14858
Qy 911 aggcgcagacgtctctctgacacaaatgtttatatacccaaccacacactgaattgt 965
Db 14859 AGGCGCAAGCTCTCCCTGAGCAAAATGTTGTATATACCAACCACTGAATTTGT 14913

RESULT 8
AC004033 145356 bp DNA PRI 31-MAY-2000
DEFINITION Homo sapiens Chromosome 22q11.2 PAC Clone p.m11 In BCR12-GGT
ACCESSION AC004033
REGION, complete sequence.
KEYWORDS HTG.
SOURCE human.

```

ORGANISM	Homo sapiens
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
AUTHORS	Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.
JOURNAL	1 (bases 1 to 145356)
REFERENCE	Budarf,M.L. and Emanuel,B.S.
AUTHORS	Unpublished
JOURNAL	2 (bases 1 to 145356)
REFERENCE	Zhang,G., Lao,V., Zhan,M. and Roe,B.A.
AUTHORS	Homo sapiens Chromosome 22q11.2 BAC Clone p_m11 In BCR12-GGT Region
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 145356)
AUTHORS	Zhang,G., Lao,V., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (26-JAN-1998) Department Of Chemistry And Biochemistry,
AUTHORS	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
JOURNAL	OK 73019, USA
REFERENCE	4 (bases 1 to 145356)
AUTHORS	Zhang,G., Lao,V., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (09-DEC-1998) Department Of Chemistry And Biochemistry,
AUTHORS	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
JOURNAL	OK 73019, USA
REFERENCE	5 (bases 1 to 145356)
AUTHORS	Zhang,G., Lao,V., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (26-JAN-1999) Department Of Chemistry And Biochemistry,
AUTHORS	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
JOURNAL	OK 73019, USA
REFERENCE	6 (bases 1 to 145356)
AUTHORS	Zhang,G., Lao,V., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (04-FEB-1999) Department Of Chemistry And Biochemistry,
AUTHORS	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
JOURNAL	OK 73019, USA
REFERENCE	7 (bases 1 to 145356)
AUTHORS	Zhang,G., Lao,V., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (09-APR-1999) Department Of Chemistry And Biochemistry,
AUTHORS	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
JOURNAL	OK 73019, USA
REFERENCE	8 (bases 1 to 145356)
AUTHORS	Zhang,G., Lao,V., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (12-APR-1999) Department Of Chemistry And Biochemistry,
AUTHORS	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
JOURNAL	OK 73019, USA
REFERENCE	9 (bases 1 to 145356)
AUTHORS	Zhang,G., Lao,V., Zhan,M., Budarf,M.L., Emanuel,B.S. and Roe,B.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (31-MAY-2000) Department Of Chemistry And Biochemistry,
AUTHORS	The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
JOURNAL	OK 73019, USA
COMMENT	On Apr 12, 1999 this sequence version replaced gi:4580479.
FEATURES	Because these overlapping clones came from different libraries
SOURCE	there are numerous instances of insertions, deletions, and single
location/Qualifiers	nucleotide polymorphisms in the overlapping regions below.
1. 145356	AC005500(p52f6) 112352 192592 (0) overlaps AC004033(p_m11) 1 78509
/organism="Homo sapiens"	(668847) AC004033(p_m11) 89227 145356 (0) overlaps AC007050(dac32) 1
/db_xref="taxon:9606"	59304 (104604) :
/chromosome="22q11.2"	Location/Qualifiers
/map="22q11.2"	1. 145356
/clone="p_m11"	
BASE COUNT	34453 a 36720 c 37358 g 36825 t
ORIGIN	
Query Match	14.4%: Score 250; DB 9; Length 145356;
Best Local Similarity	100.0%; Pred. No. 3.8e-112;

Matches	250; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY 716	agcccccagcaacaattcagcagccacgcatgacgacacagctccgcctcccaagg							775
Db 103886	AGGCCAGCCAGCCAAATTTCAGAGCCAGCCAGGATGACAGCACCAGAGCTTCGCCAGG							103945
QY 776	ctctgccccagcagctgtgacgagatgatacaacacagcaccaagccgcgcacacagc							835
Db 103946	CTCTGCCCCAGCAGTGTGAGGAGATGATCATCACACAGCACCACCAAGCCGCACACAGC							104005
QY 836	ccagcagcctcagtlgtctcagaacaacacatcaccaactccgcgcacagtcgacacc							895
Db 104006	CCAGCAGACCTTCAGATTCTCAGAAACCAACATCATCAACTCCCGCACAGTCCAGACCC							104065
QY 896	agccttggtgtcacagagcgcaagctctccctggaacaattgttatataccaacacac							955
Db 104066	AGCCTTGGTGTTCACAGCGCAGAGCTCTCCCTGGCAAAATGTTGTATACCAACACGAC							104125
QY 956	tgaatttgt 965							
Db 104126	TGAATTGTGT 104135							
RESULT 9								
AK000003	AK000003	4701 bp	mRNA	PRI	25-Apr-2000			
LOCUS	Homo sapiens mRNA for FLJ00003 protein, partial cds.							
DEFINITION	AK000003							
ACCESSION	AK000003							
VERSION	AK000003.1 GI:7209306							
KEYWORDS	FLJ00003 protein; fls (full insert sequence).							
SOURCE	Homo sapiens adult spleen cDNA to mRNA, clone:as00003.							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.							
AUTHORS	Ohara,O., Kikuno,R., Nagase,T. and Okumura,K.							
TITLE	1 (bases 1 to 4701)							
COMMENT	The nucleotide sequence of a long cDNA clone isolated from human spleen							
JOURNAL	Published Only in Database (2000) In press							
REFERENCE	2 (bases 1 to 4701)							
AUTHORS	Ohara,O., Kikuno,R., Nagase,T. and Okumura,K.							
TITLE	Direct Submission							
JOURNAL	Submitted (14-FEB-2000) to the DDBJ/EMBL/GenBank databases. Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdna@kazusa.or.jp, URL:http://www.kazusa.or.jp, Tel:81-438-52-3913, Fax:81-438-52-3914)							
COMMENT	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology: cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection Kazusa DNA Research Institute.							
FEATURES	Location/Qualifiers							
source	1..4701							
	/organism="Homo sapiens"							
	/db_xref="taxon:9606"							
	/clone="as00003"							
	/dev_stage="adult"							
	/tissue_type="spleen"							
	/note="vector: pBluescriptII SK plus"							
	3081..3788							
gene	/gene="FLJ00003"							
	<3081..3788							
CDS	/gene="FLJ00003"							
	/note="For this clone, Genemark analysis triggered an alert for spurious CDS split. The result of Genemark analysis of the nucleotide sequence of this clone is accessible through http://www.kazusa.or.jp.							
	Start codon is not identified."							
	/codon_start=1							
	/product="FLJ00003 protein"							
	/protein_id="BAA92228.1"							
	/db_xref="GI:7209307"							

	Query Match	7.6%;	Score 132;	DB 6;	length 386;
	Best Local Similarity	99.1%;	Pred. NO. 9.3e-54;		
	Matches 233;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps	0
QY	629	agcagcgcgcgaatagcacacagctcgcagctcccaacaacagcaacagcagcagcagc	688		

RESULT	13
AC005500	
LOCUS	AC005500 192592 bp DNA PRI 31-MAY-2000
DEFINITION	Homo sapiens 22q11 PAC Clone Clone p52t6 In DGR Region, compl
ACCESSION	sequence. AC005500

VERSION AC005500.2 31:7798766
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Homo sapiens Chromosome 22q11 PAC Clone p52f6 In DSCR Region
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (20-AUG-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 3 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (13-SEP-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 4 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (23-SEP-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 5 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (01-OCT-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 6 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (01-OCT-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 7 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (14-OCT-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 8 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (23-OCT-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 9 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (30-OCT-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 10 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (03-NOV-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 11 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-1998) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 12 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission

JOURNAL Submitted (13-MAY-2000) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE 13 (bases 1 to 192592)
 AUTHORS Chen, F., D.L., Do, T., Dumanski, J.P. and Roe, B.A.
 TITLE Direct Submission
 JOURNAL Submitted (31-MAY-2000) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 COMMENT On May 13, 2000 this sequence version replaced gi:3900817.
 Because these overlapping clones came from different libraries
 there are numerous instances of insertions, deletions, and single
 nucleotide polymorphisms in the overlapping regions below.
 AC00773(b562f10) 5026 182617 (0) overlaps AC005500(p52f6) 1 177555
 (15037) AC005500(p52f6) 112352 192592 (0) overlaps AC004033(p_m11)
 1 78509 (66847).
 FEATURES
 location/Qualifiers
 1..192592
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="22q11"
 /clone="p52f6"
 BASE COUNT 47154 a 49045 c 48565 g 47828 t
 ORIGIN
 Query Match 4.5%; Score 79; DB 9; Length 192592;
 Best Local Similarity 100.0%; Pred. No. 6.4e-28;
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 atgaggaaagctggtggtgacacagtaaatccagcagaagatagagagcatgttc 60
 Db 188530 ATGAGGAAGCTGCTGTGGCAGACAGTAATCCAGCAGATATGGAGACCATGTTTC 188589
 Oy 61 ctgaagcgaagcccgag 79
 Db 188590 CTGAGGCCAGACCCGAG 188608
 RESULT 14
 AF125313 107642 bp DNA ROD 22-MAR-2000
 LOCUS Mus musculus x chromosome: Cg1/Xap80 gene, complete sequence.
 DEFINITION AF125313
 ACCESSION AF125313.1 GI:6650998
 VERSION HTG.
 KEYWORDS house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 107642)
 AUTHORS Platzer, M., Zhao, W., Herman, G.E. and Rosenthal, A.
 TITLE Comparative sequence analysis of the mouse Mtm locus and the
 corresponding region of human Xq28
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 107642)
 AUTHORS Platzer, M.
 TITLE Direct Submission
 JOURNAL Submitted (02-FEB-1999) Genome Analysis, Institute of Molecular
 Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
 COMMENT This sequence is part of a larger genomic contig. The start of this
 sequence is directed towards the centromere. The end
 (107393..107642) of this sequence overlaps with the start of the
 neighboring sequence deposited in Genbank Accession Number
 AF125314.
 FEATURES
 location/Qualifiers
 1..107642
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="X"
 /clone="508014"
 /clone_1ib="ClbC07 [129/Sv]"
 repeat_region 607..645

```

repeat_region /rpt_family="AT_rich"
complement(1452..1611)
/rpt_family="L1MC4"
repeat_region 1789..2073
/rpt_family="M1T1A1"
complement(2164..2898)
/rpt_family="L1MC4"
repeat_region /rpt_family="L1MC4"
complement(3067..3238)
/rpt_family="L1MD3"
repeat_region /rpt_family="L1MC4"
complement(3475..3580)
/rpt_family="L1MC4"
repeat_region /rpt_family="L1MC4"
complement(3668..3720)
/rpt_family="CA)n"
repeat_region /rpt_family="L1MC4"
complement(4236..4469)
/rpt_family="L1ME"
repeat_region 4539..4567
/rpt_family="AT_rich"
complement(4653..4710)
/rpt_family="B4"
repeat_region /rpt_family="B4"
complement(4922..4965)
/rpt_family="A-rich"
repeat_region /rpt_family="L1M4"
complement(5212..5684)
/rpt_family="L1M4"
repeat_region /rpt_family="L1M4"
complement(5769..5799)
/rpt_family="CA)n"
repeat_region /rpt_family="L1M4"
complement(5851..5986)
/rpt_family="RSINE1"
repeat_region /rpt_family="L1M4"
complement(6012..6168)
/rpt_family="L1M4"
repeat_region /rpt_family="L1M4"
complement(6189..6305)
/rpt_family="L1M4"
repeat_region /rpt_family="L1M4"
complement(6306..6350)
/rpt_family="CA)n"
repeat_region /rpt_family="L1M4"
complement(6351..6409)
/rpt_family="L1M4"
repeat_region /rpt_family="L1M4"
complement(6444..6621)
/rpt_family="B3"
repeat_region /rpt_family="B3"
complement(6647..6698)
/rpt_family="B1-F"
repeat_region /rpt_family="B1-F"
complement(6730..6884)
/rpt_family="L1M4"
repeat_region 7032..7157
/rpt_family="B1-MM"
repeat_region 7158..7186
/rpt_family="B1-MM"
complement(7437..7635)
/rpt_family="B3"
repeat_region /rpt_family="B3"
complement(7665..7723)
/rpt_family="L1M4"
repeat_region 7768..8013
/rpt_family="Lx2"
repeat_region 8021..8097
/rpt_family="AT_rich"
complement(8117..8208)
/rpt_family="RSINE1"
repeat_region /rpt_family="RSINE1"
complement(8255..8430)
/rpt_family="B4A"
repeat_region /rpt_family="B4A"
complement(8501..8860)
/rpt_family="L1M4"
repeat_region /rpt_family="L1M4"
complement(8900..9318)
/rpt_family="Lx8"
repeat_region /rpt_family="Lx8"
complement(9682..9946)
/rpt_family="L1MEC"
repeat_region /rpt_family="L1MEC"
complement(10091..10619)
/rpt_family="L1MD"
repeat_region /rpt_family="L1MD"
complement(11193..11366)
/rpt_family="L1MC1"
repeat_region 11510..11652
/rpt_family="B1-MM"
complement(11655..12179)
/rpt_family="L1MC2"
repeat_region /rpt_family="L1MC2"
complement(12684..12771)
/rpt_family="L1MB2"
repeat_region 14236..14266
/rpt_family="POLY_A"
repeat_region 15218..15552
/rpt_family="M1T1C"
complement(15570..15798)
/rpt_family="L1M2"
repeat_region /rpt_family="L1M2"
complement(15898..16075)
/rpt_family="L1M2"
repeat_region /rpt_family="L1M2"
complement(16410..16698)
/rpt_family="L1M1"
repeat_region /rpt_family="L1M1"
complement(17963..18084)
/rpt_family="MIR"
repeat_region /rpt_family="MIR"
complement(19022..19439)
/rpt_family="Charlie4a"
repeat_region /rpt_family="Charlie4a"
complement(19936..20069)
/rpt_family="MRE63"
repeat_region /rpt_family="MRE63"
complement(21355..21462)
/rpt_family="CA)n"
repeat_region /rpt_family="CA)n"
complement(21475..21592)
/rpt_family="B1-F"
repeat_region /rpt_family="B1-F"
complement(22549..22708)
/rpt_family="MIR"
repeat_region /rpt_family="MIR"
complement(22835..22933)
/rpt_family="URR1A"
repeat_region /rpt_family="URR1A"
complement(23475..23603)
/rpt_family="B4A"
repeat_region /rpt_family="B4A"
complement(23609..23749)
/rpt_family="B1-MM"
repeat_region /rpt_family="B1-MM"
complement(23761..23848)
/rpt_family="B4A"
repeat_region 25114..25177
/rpt_family="TA)n"
complement(25206..25436)
/rpt_family="Lx8"
repeat_region /rpt_family="Lx8"
complement(25764..26095)
/rpt_family="Lx9"
repeat_region /rpt_family="Lx9"
complement(26161..26220)
/rpt_family="GA)n"
repeat_region /rpt_family="GA)n"
complement(26225..26249)
/rpt_family="CA)n"
repeat_region 26630..26742
/rpt_family="RSINE1"
complement(26743..26786)
/rpt_family="CA)n"
repeat_region 26791..26819
/rpt_family="CA)n"
complement(30815..30869)
/rpt_family="CA)n"
repeat_region 30872..31065
/rpt_family="L1MB8"
repeat_region /rpt_family="B3"
complement(31078..31214)
/rpt_family="B1-F"
repeat_region /rpt_family="B1-F"
complement(31246..31270)
/rpt_family="Lx9"
repeat_region /rpt_family="Lx9"
complement(31271..31293)
/rpt_family="CA)n"
repeat_region /rpt_family="CA)n"
complement(31294..31517)
/rpt_family="Lx9"
repeat_region /rpt_family="Lx9"
complement(31518..31780)
/rpt_family="polypurine"
repeat_region /rpt_family="polypurine"
complement(32109..32186)
/rpt_family="B1-F"
repeat_region 32193..32380
/rpt_family="L1-MM"
repeat_region 32381..32410
/rpt_family="TAA)n"
complement(32470..32548)
/rpt_family="ID3"
repeat_region /rpt_family="ID3"
complement(32877..33115)
/rpt_family="URR1B"
repeat_region 34310..34331
/rpt_family="TA)n"
complement(34371..34548)

```


Wed May 1 16:36:04 2002

us-09-668-119-1.oli.rge

Page 15

1
2
3
4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 30, 2002, 19:49:19 ; Search time 182.29 Seconds

(without alignments)
8183.367 Million cell updates/sec

Title: US-09-668-119-1
Perfect score: 1740
Sequence: 1 atgaggaagtggtgtggtc.....cacatgcctccagcatga 1740

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N.Geneseq-1101:*

- 1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1518	87.2	2556	22	AA159455 Human polynucleoti
2	848	48.7	3458	22	AA161241 Human polynucleoti
3	777	44.7	1388	21	AA159039 Human secreted pro
4	239	13.7	316	22	AA120017 Probe #9950 for ge
5	239	13.7	316	22	AA145215 Probe #13901 used
6	239	13.7	316	22	AA105725 Probe #5716 used t
7	239	13.7	436	22	AA110775 Probe #708 for gen
8	239	13.7	436	22	AA132033 Probe #719 used to
9	239	13.7	436	22	AA100701 Probe #692 used to
10	132	7.6	386	22	AA165986 Novel human polynu
11	73	4.2	232	21	AA109371 Human secreted pro

12	45	2.6	203	19	AAV30271
13	45	2.6	1310	20	AAV17226
14	45	2.6	3263	20	AAV08558
15	45	2.6	3263	20	AAV08560
16	45	2.6	4286	20	AAV08559
17	44	2.5	3217	10	AAV91578
18	44	2.5	3217	12	AAQ12002
19	44	2.5	4180	10	AAV91773
20	42	2.4	293	22	AAI21460
21	42	2.4	293	22	AAI46749
22	42	2.4	293	22	AAI07156
23	42	2.4	397	20	AAV89891
24	42	2.4	459	22	AAI12263
25	42	2.4	459	22	AAI33617
26	42	2.4	459	22	AAI02177
27	42	2.4	1776	20	AAZ23425
28	42	2.4	2415	18	AAV86757
29	42	2.4	2453	18	AAV86758
30	41	2.4	1853	19	AAV06751
31	41	2.4	4279	20	AAZ23430
32	41	2.4	4789	20	AAV80992
33	40	2.3	107	21	AAV18151
34	39	2.2	2171	21	AAV39075
35	39	2.2	2224	19	AAV35472
36	39	2.2	2733	20	AAV84442
37	39	2.2	8564	21	AAV4783
38	38	2.2	102	22	AAI24904
39	38	2.2	102	22	AAI50516
40	38	2.2	374	22	AAV66719
41	38	2.2	491	22	AAI15681
42	38	2.2	491	22	AAI37333
43	38	2.2	1635	21	AAV62682
44	38	2.2	2333	22	AAV75339
45	38	2.2	3302	22	AAI58588

ALIGNMENTS

RESULT 1	
ID	AA159455 standard; cDNA: 2556 BP.
AA159455	
AC	AA159455;
XX	
DT	22-OCT-2001 (first entry)
XX	
DE	Human polynucleotide SEQ ID NO 1658.
XX	
KW	Human; nootropic; immunosuppressant; cytosolic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW	leukaemia; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000WO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0695036.
PR	29-NOV-2000; 2000US-0727344.
XX	

Glutamine rich reg
SCA2 gene CAG repe
Transgene for epit
Transgene for epit
Transgene for epit
Rat androgen recep
Full-length rat an
Rat androgen recep
Probe #11393 for g
Probe #15435 used
Probe #7147 used t
Sphincerebellar at
Probe #2196 for ge
Probe #2303 used t
Probe #2168 used t
Human Machado-Jose
CDNA of the M3/6 g
CDNA of the M3/6 g
Dentatorubral-pall
Human atrophin I D
Human steroid rece
Human secreted pro
Human secreted pro
Rat RSK3 coding se
Mouse brain CNG-1
Human OREF ORF38
Probe #14837 for g
Probe #19202 used
Novel human polynu
Probe #5614 for ge
Probe #6019 used t
Human ASH1 gene.
Human TGF-beta rec
Human polynucleoti

Db 1794 atcaacaagatcgaacagaacagacagaaaagacctgagtaagatgaagacgtt 1853
 QY 1510 ctgagacattctgacagacccctcgaaagcggtgtcccttgaaagccttgcacaaagtgtgag 1569
 Db 1854 ctgagacattctgacagacccctcgaaagcggtgtcccttgaaagccttgcacaaagtgtgag 1913
 QY 1570 atcgcccttgagaaacacaaagatgacat-gcggtgtccacaccccccaccccggtgtg 1628
 Db 1914 atcgcccttgagaaacacaaagatgacatgaggtgtccacaccccccaccccggtgtg 1973
 QY 1529 ccaccgacacaaacagagatgacatgacagcgctccttgatgacgctcctgcgcacacatc 1688
 Db 1974 ccaccgacacaaacagagatgacatgacagcgctccttgatgacgctcctgcgcacacatc 2033
 QY 1689 cgctcactgtcttaacacattccctgtacacgacattcgttcagacatga 1740
 Db 2034 cgctcactgtcttaacacattccctgtacacgacattcgttcagacatga 2085

RESULT 3
 AAC59039 standard; cDNA; 1388 BP.
 AAC59039;
 02-FEB-2001 (first entry)
 Human secreted protein coding sequence SEQ ID NO: 60.

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
 KW antiviral; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW candidate; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological diseases; infection; human; secreted protein; ss.

OS Homo sapiens.
 FN W0200055175-A1.
 PD 21-SEP-2000.
 PF 09-MAR-2000; 2000WO-US06049.
 PR 12-MAR-1999; 99US-0124144.
 PR 11-JUN-1999; 99US-0138574.
 PR 03-DEC-1999; 99US-0168667.
 XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;
 DR WPI: 2000-638175/61.
 DR P-PSDB; AAB27609.

PT Isolated nucleic acid molecule encoding a human secreted protein is
 used in preventing, treating or ameliorating a medical condition -
 Claim 1; Page 372; 428pp; English.

The invention relates to the isolation of genes AAB27560-B27609 encoding
 30 human secreted proteins AAB27560-B27609. The genes can be used to
 generate fusion proteins by linking to the gene for the human
 immunoglobulin G Fc portion (SEQID1) for increasing the stability of
 the fusion protein as compared to the human protein only. The genes and
 proteins are useful for preventing, ameliorating or treating medical
 conditions, e.g. by protein or gene therapy. The genes are isolated
 from a range of human tissues disclosed in the specification. The
 nucleic acids, proteins, antibodies and (ant)agonists are useful in
 the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 and ovarian cancer, and other cancers of the adrenal gland, bone, bone
 marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.

SQ Sequence 1388 BP; 364 A; 378 C; 343 G; 297 T; 6 other.

Query Match 44.7%; Score 777; DB 21; Length 1388;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 927; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 25 agtaaatccagaaagatgtagagagccatgtttcctgaagagccaaagaccggagaa 84
 Db 122 agtaaatccagaaagatgtagagagccatgtttcctgaagagccaaagaccggagaa 181
 QY 85 tactttcctcgttgagcagagctcattatccatttgcagacattcataaagaatct 144
 Db 182 tactttcctcgttgagcagagctcattatccatttgcagacattcataaagaatct 241
 QY 145 caagcttcgctcagatgataatcctatgatacctccagagcctgagcgtctgag 204
 Db 242 caagcttcgctcagatgataatcctatgatacctccagagcctgagcgtctgag 301
 QY 205 ggaagccgctggaattgcatgctcctcgcgggcccgggagacagttcttgaggagtgag 264
 Db 302 ggaagccgctggaattgcatgctcctcgcgggcccgggagacagttcttgaggagtgag 361
 QY 265 agctttgtgcatgtagagaaagccaatgtctcctcctcagaggcagccgctctctgagcccg 324
 Db 362 agctttgtgcatgtagagaaagccaatgtctcctcctcagaggcagccgctctctgagcccg 421
 QY 325 ggaatggccctcagagcagatgctgtcgtctcagcagcaatccacagaccagacgtgag 384
 Db 422 ggaatggccctcagagcagatgctgtcgtctcagcagcaatccacagaccagacgtgag 481
 QY 385 ctccagcagatgtagcgtctgtagcagcagcagcagcagcagcagcagcagcagcagcag 444
 Db 482 ctccagcagatgtagcgtctgtagcagcagcagcagcagcagcagcagcagcagcagcag 541
 QY 445 ggcgcgctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 504
 Db 542 ggcgcgctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 601
 QY 505 ggcgcgctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 564
 Db 602 ggcgcgctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 661
 QY 565 cag 624
 Db 662 cag 721
 QY 625 caacag 684
 Db 722 caacag 781
 QY 685 cag 744
 Db 782 cag 841
 QY 745 atgcag 804
 Db 842 atgcag 901
 QY 805 cacacacag 864
 Db 902 cacacacag 961
 QY 865 ccatacaaatcccgccacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 924
 Db 962 ccatacaaatcccgccacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1021

1
2
3
4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 30, 2002, 16:38:07 ; Search time 1518.73 Seconds
(without alignments)
12311.375 Million cell updates/sec

Title: US-09-668-119-1
Perfect score: 1740
Sequence: 1 atgaggaagctgtgtgtgc.....cacatcgttcacgacatga 1740

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size : 0
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
EST:
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estipl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_estl:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rnd:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	566	32.5	730	11	BG766214 602737942
2	551	31.7	2752	12	BC005027 Homo sapi
3	509	29.3	620	10	BC384305 601272684
4	475	27.3	526	11	BE019568 ba83f05.y
5	450	25.9	776	11	BG750043 602708813
6	431	24.8	482	11	AA664125 ac06c04.s
7	418	24.0	914	11	BG468341 602509947
8	399	22.9	756	11	BG746380 602703648
9	395	22.7	886	11	BG281923 602403104
10	385	22.1	723	11	BE279307 601157640
11	357	20.5	486	10	AW747905 QV1-BN000
12	347	19.9	757	10	AL046886 DKFZP586E

C	13	340	19.5	501	10	BE163519	BE163519 QV3-HT046
	14	327	18.8	976	11	BG122694	BG122694 602351822
	15	319	18.3	858	11	BG386117	BG386117 602455292
C	16	308	17.7	524	10	BE410770	BE410770 601301681
	17	301	17.3	498	10	AM603679	AM603679 CM0-CN003
	18	292	16.8	656	10	BE390284	BE390284 601285167
	19	287	16.5	557	11	BG421332	BG421332 602451441
	20	283	16.3	1019	11	BG352275	BG352275 602403789
	21	267	15.3	424	10	AM501171	AM501171 UI-HF-BP0
	22	267	15.3	428	11	BE889171	BE889171 601513434
	23	260	14.9	386	11	BG982754	BG982754 PM0-CN015
	24	255	14.7	1157	11	BG177319	BG177319 602313922
	25	254	14.6	393	10	AM499614	AM499614 UI-HF-BP0
	26	252	14.5	391	10	AM500833	AM500833 UI-HF-BP0
	27	242	13.9	313	11	N49344	N49344 Y236e10.r1
	28	241	13.9	386	11	BG982761	BG982761 PM0-CN015
	29	236	13.6	434	11	BF346622	BF346622 602021224
	30	216	12.4	942	11	BG338685	BG338685 602436442
C	31	211	12.1	335	11	B1011820	B1011820 PM1-EN006
	32	185	10.6	939	11	BG336790	BG336790 602405533
	33	181	10.4	235	11	BF994170	BF994170 CM2-CN016
	34	172	9.9	733	10	BE391135	BE391135 601286095
	35	170	9.8	348	11	T07163	T07163 EST05052.Fe
C	36	169	9.7	258	11	BF755836	BF755836 CM3-CT057
	37	164	9.4	810	11	BG686794	BG686794 602650765
	38	164	9.4	1043	11	BF974180	BF974180 602243806
C	39	158	9.1	443	11	B1024741	B1024741 PM0-MT020
	40	143	8.2	668	11	BG766083	BG766083 602738467
	41	138	7.9	255	11	BF870481	BF870481 IL3-ET011
C	42	136	7.8	568	11	B1089323	B1089323 602853282
	43	135	7.8	385	11	R17502	R17502 Y902d12.r1
C	44	130	7.5	200	11	BF084780	BF084780 IL3-CT053
	45	121	7.0	301	10	AA069235	AA069235 zml1e01.r

ALIGNMENTS

RESULT 1
LOCUS BG766214 730 bp mRNA
DEFINITION 602737942F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4862938 5', mRNA sequence.
ACCESSION BG766214
VERSION BG766214.1 GI:14076867
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 730)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcaps-remail.nih.gov
Tissue Procurement: ATCC/DCFT/DP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LCM1724 row: d column: 11
High quality sequence stop: 716.
location/Qualifiers
1..730
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4862938"
/clone_id="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"

[illegible]

Db	422	TCCAGGCTCAGCAGAGTGGCCATGCACACAGCAAGTTCCAMGACGTAGTGACAGCGACAGC	363
OY	548	agctccgagcgacgacgacgacgacgacgacatctaattgaattgatcatcaaatcagc	607
Db	362	AGCTTCAGCAGCAGCACACACACACACAGCATTAATTAAATTGATCATCAATAATTCAGC	303
OY	608	aacagatacacgacgacgacaacgacgactgcagtcgaaatagaacagctgcagctccaacac	667
Db	302	AACGAGATATACGACGACACACACACACTGACGGAATTAGCACAGCTGCAGCTCCAACAC	243
OY	668	agcaaacgacgacgacgacgacgacgacgacgacgacgacgttttggaggccccagccac	727
Db	242	AGCAAACGACGACGACACACACACACACACAGCAGCAGCAGCAGCTTTGCAAGGCCACGAC	183
OY	728	caatttcagcgaccacgaatgcagcagcgcacagctccgccctcccagagctctgcgccagc	787
Db	182	CAATTTCAGCAGCCACCGAATGACGACGACCAGCTCCGCCCTCCAGAGCTCTGCCCCAGC	123
OY	788	agctcagcagatgcatcacacacagcaaccaaccagcgcgcacacacagccccagcagctc	847
Db	122	AGTGCGACGACGATGCATCACACACAGCACACACAGCCGCCACACAGCCCCAGCAGCTTC	63
OY	848	cagttgcttgaaaccaaaccctcgaacctccgcgcacagctcgagaccagccttgtgagt	907
Db	62	CAGTTGCTCGAACAACCAATCACTCACTCCGCAACAGTAGTGACAGCCACAGCTTTGGTGT	3
OY	908	ca 909	
Db	2	CA 1	
RESULT	7		
LOCUS	BG468341	914 bp mRNA	EST 21-MAR-2001
DEFINITION	602509947F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4644299 5'		
KEYWORDS	BG468341	mRNA sequence.	
SOURCE	BG468341.1	GI:13400611	
ORGANISM	Homo sapiens	human.	
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/		
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: rgs@rsf-research.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: NIH Intramural Sequencing Center Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LCM415 row: n column: 12 High quality sequence stop: 812. Location/Qualifiers 1. .914 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4644299" /clone_lib="NIH_MGC_15" /tissue_type="adenocarcinoma cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: colon; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(C). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using Zap-cDNA synthesis kit"		

BASE COUNT (Stratagene) and Superscript II RT (Life Technologies)"
 ORIGIN 225 a 305 c 229 g 155 t

Query Match 24.0%; Score 418; DB 11; Length 914;
 Best Local Similarity 99.8%; Pred. No. 2.9e-181;
 Matches 538; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1203 caagtcagctctgcccctcccccattccagtagcttccctgcccagccctcaccgca 1262
 DB 3 CAACCTCAGCTCTGACCTCTGCCCCATCTCCAGTAGCTTCTGCCCCACCTCACC 62
 OY 1263 gccctccagagccagagagagagagagagagagagagagagagagagagag 1322
 DB 63 GCCCTCCAGAGCCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122
 OY 1323 acccttaaacacacctgtggaacccagctctgtcatgagccagctgtgctcagc 1382
 DB 123 ACCTTTAAACACACCTGTGAACCCAGCTGTGTATGAGCCAGCTGGCTCAGCAG 182
 OY 1383 tgaag 1442
 DB 183 TGAG 242
 OY 1443 ccgcatgatcaacagatcgacagagagagagagagagagagagagagagagag 1502
 DB 243 CCGCATGATCAACAGATCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 302
 OY 1503 gaggctcttgagacattctgacagagagagagagagagagagagagagagagag 1562
 DB 303 GAGCCTTTGAGCAATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362
 OY 1563 gtgtgagatcgcccttgagagagagagagagagagagagagagagagagagag 1621
 DB 363 GTGTGAGATCGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 422
 OY 1622 ccgagtgacag 1681
 DB 423 CCGGCTGACAG 482
 OY 1682 caagctcgctcagctctgctcagagagagagagagagagagagagagagag 1740
 DB 483 CAACATCGCTCAGCTGTCTTCAACAGATCTCCCTGACAGAGAGAGAGAGAG 541

RESULT 8
 LOCUS BG746380 756 bp mRNA EST 15-MAY-2001
 DEFINITION 602703648P1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4857083 5',
 mRNA sequence.
 ACCESSION BG746380
 VERSION BG746380.1 GI:14057033
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 756)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@biml.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: NIH Intramural Sequencing Center (LNL)
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM1708 row: P column: 12
 High quality sequence stop: 756.

FEATURES
 source

Location/Qualifiers
 1. 756
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4857083"
 /clone_1b="NIH_MGC_15"
 /tissue="adenoecarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: colon; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(9). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)"
 BASE COUNT 179 a 264 c 186 g 126 t 1 others

Query Match 22.9%; Score 399; DB 11; Length 756;
 Best Local Similarity 99.8%; Pred. No. 1.6e-172;
 Matches 519; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1222 gccacatctccagtagcttctgcccagccctcaccgagccctccagagccagtg 1281
 DB 17 GCCCATCTCCAGTAGAGCTTCCGCCAGCCCTCACCCAGCCCTCCAGAGCCAGTG 76
 OY 1282 acgagcgag 1341
 DB 77 ACGGCGGAG 136
 OY 1342 aacccagctctgcatgagagagagagagagagagagagagagagagagagag 1401
 DB 137 AACCCAGCTCTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196
 OY 1402 gacaaagtgaag 1461
 DB 197 GACAACTGAAG 256
 OY 1462 gacaaag 1521
 DB 257 GACAAAG 316
 OY 1522 acagagccctcgaagcgtgtgcccctggaagagagagagagagagagagag 1581
 DB 317 ACAGAGCCCTCGAAGCGGTGCTCCCTGAAGAGAGAGAGAGAGAGAGAGAGAG 376
 OY 1582 aaactcaagatgacat-gcgtgtgcccactcccccagccggtgagcagagca 1640
 DB 377 AAACCTCAAGATGACATGTGGCGGGCCACTCCGCCAGCCGCCGCTGACCA 436
 OY 1641 caagcagtaacctatgacagagagagagagagagagagagagagagagagagag 1700
 DB 437 CACAGTACTATGATGACAGCGGCTCTGTGATGCGGTCTGCAACATCGCTACCTGTC 496
 OY 1701 ttcaaccattccctgtacagagagagagagagagagagagagagagagagagag 1740
 DB 497 TTCAACATTCCCTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 536

RESULT 9
 LOCUS BG281923 886 bp mRNA EST 21-FEB-2001
 DEFINITION 602403104P1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4545433 5',
 mRNA sequence.
 ACCESSION BG281923
 VERSION BG281923.1 GI:13030849
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 886)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCMD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: LNCMI230 row: g column: 02
High quality sequence stop: 876.
Location/Qualifiers
1. 886
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4545433"
/clone_lib="NIH-MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 234 a 299 c 231 g 122 t
ORIGIN

Query Match 22.7%; Score 395; DB 11; Length 886;
Best Local Similarity 100.0%; Pred. No. 1.1e-170;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 304 cagcgcctctctggaagctggagatggccctacacagatgctgtgtctacgca 363
DB 15 CAGCCGCTCTCTGGACCTGGGGATGGCCCTCACAGCATGCTCTGTACGGCA 74
QY 364 acccacaagaccagctcagctcagcagctgctcagcagcagcagcagcagc 423
DB 75 ACTCCACAGACCCAGCTGCACTCCAGCAGGTGGCTGAGCAGCAGCAGCAG 134
QY 424 cagttcagaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 483
DB 135 CAGTTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 194
QY 484 cagttcagaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 543
DB 195 CAGTTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 254
QY 544 cagcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 603
DB 255 CAGCAGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 314
QY 604 cagcaacaagatatacagcagcagcagcagcagcagcagcagcagcagcagc 663
DB 315 CAGCAACAGATACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 374
QY 664 caacaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 698
DB 375 CAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 409
RESULT 10
BE279307 723 bp mRNA EST 13-JUL-2000
LOCUS 601157640F1 NIH-MGC_21 Homo sapiens CDNA clone IMAGE:3504608 5',
DEFINITION mRNA sequence.
ACCESSION BE279307

VERSION BE279307.1 GI:9154297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 723)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: LNCMI79 row: o column: 09
High quality sequence stop: 661.
Location/Qualifiers
1. 723
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3504608"
/clone_lib="NIH-MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 177 a 208 c 224 g 114 t
ORIGIN

Query Match 22.1%; Score 385; DB 10; Length 723;
Best Local Similarity 99.6%; Pred. No. 4.2e-166;
Matches 485; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 atgaggaagctggtgtgtgacacagtaataccagcaagataggaagcagtttc 60
DB 107 ATGAGGAAGCTGTTGTGGCACAGTAATCCAGCAAGATATGAGAGGCAATGTTTC 166
QY 61 ctgaagggccaagaccgggaagaaacttctctcgtggccaagctcatatcatatt 120
DB 167 CTGAAGGCCAAGACCCGGAGCAATACCTTTCTCTGTCGTCAGGCTCATATTCATTTT 226
QY 121 cgaagcatcataaagaanaalctcaagcttcgctcagtgatcatatgaatgcaactccag 180
DB 227 CGAGCATTCATTAACAAGAAATCTCAAGCTCCGTCAGATGATCTAATGAACACTCCAG 286
QY 181 agcctgactgcygagcctgctgcygagagcgtgtaattgcatatgctccctcggggcccg 240
DB 287 AACTCTGACTGGGGAGCTGCTGCGGAGCCGCTGGAATTGGCATGCCCTCTCGGGGCCG 346
QY 241 ggaagctctctggggagatgtagcttggtagcagtgagagcagcaatgtctctca 300
DB 347 GGACAGTCTCTTGCGGGGATGGCTTGCTTGCTGACAGGAGCAATGTCTCTCA 406
QY 301 gggcagccgctctctggaagcctcgggagatggccctcagacatgagctgtgtctacg 360
DB 407 GGGCAGCCGCTCTCTGGGAGCTCGGGGATGGCCCTCAAGAGATGGCTGTGTCTACG 466
QY 361 gcaactcacaagaccagctgagctcagcagcagctgagcagcagcagcagcagc 420
DB 467 GCAACTCCACAGACCCAGCTGACGCTCCAGCAGGTGGGGCTGCACACACACAGCAACAG 526
QY 421 cagcagttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480

Db 123 CGAGACATTCATACAGAAATCTCAAGCTTCCGTCAGTATCTATGAATGACCTCCAG 182

QY 181 agcctgactgaggagactgctgaggaagccgctggaattgacatgctctctcgggccg 240

Db 183 AGCCTGACTGGGAGACTCTGCTGGGAGCCGCTGGAATGGCATGCTCTCTCTGGGCCCCG 242

QY 241 ggaacatctctggggaggtgtagcttggtgcatatggagcaagcaatgctctctca 300

Db 243 GGACATCTCTGTGGGGGATGGGTACCTTGGTGGCATGGAGCAAGCAATGTCCTCTCA 302

QY 301 gggacagccctctctgagacctcgagatgagccctcacagatggtctgctctacg 360

Db 303 GGGACACCCCTCTCTGGAGCTCTGGGAGTGGCCCTCCACAGCATGCTCTCTGTCTACG 362

QY 361 gcaactcacagagaccagctgacgtcagctcagcaggtgacg 398

Db 363 GCAACTCCACAGACCAGCTGACAGTCCAGCAGGTGGC 400

RESULT 13
BE163519 501 bp mRNA EST 21-JUN-2000
LOCUS BE163519 QV3-HT0460-230200-101-b02 HT0460 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE163519
ACCESSION BE163519
VERSION BE163519.1 GI:8626240
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 501)

REFERENCE
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV3-HT0460-230200-101-b02&ts=2000-02-23&tl=1)
Seq primer: puc 18 forward
High quality sequence start: 8
High quality sequence stop: 499.
Location/Qualifiers
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0460"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 72 a 121 c 192 g 116 t
ORIGIN

Query Match 19.5%; Score 340; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 2e-145;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 735 gcaagcacagatgacagcagcagcagcctccgctccagagctctgcccagcagctgca 794

Db 453 GCAGCCACCGATGACAGCAGCAGCAGCCTCCGCTCCAGGCTGTGCCCCAGCAGTGCA 394

QY 795 gcaagatgacacacacagcagcagcagcagcagcagcagcagcagcagcagcagcag 854

Db 393 GCAGATGATGATAC 334

QY 855 tcagagcaaacacatcac 914

Db 333 TCAGAACCAACCATACACTCCGCCACAGTCCAGACCCAGCTTTGGTGTACAGGC 274

QY 915 gcaagctctccctggacaaatgltglataccacacacacacacacacacacacacacac 974

Db 273 GCAAGCTCTCCCTGGACAAATGTTGTATATACCAACACACTGAATTTGTCCGAGCTCC 214

QY 975 gatgtgtgtgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1034

Db 213 GATGTGTGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 154

QY 1035 gacagctcagcgtgccagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1074

Db 153 GACAGCTCAGGCTGCCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 114

RESULT 14
LOCUS BG122694 956 bp mRNA EST 30-JAN-2001
DEFINITION BG122694 60235182251 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4450300 5',
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RNA sequence.
ACCESSION BG122694
VERSION BG122694.1 GI:12616203
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 956)

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10235 row: k column: 05
High quality sequence stop: 611.
Location/Qualifiers
1..956
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4450300"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC library."

BASE COUNT 246 a 325 c 251 g 134 t
ORIGIN

Wed May 1 16:36:06 2002

us-09-668-119-1.oli.rst

Page 11

11
12
13
14

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 30, 2002, 17:43:22 ; Search time 92.24 Seconds
(without adjustment)

4272.243 Million cell updates/sec

Title: US-09-668-119-1

Sequence: 1 atgaggaagctgtgtgac.....cacattcgttccaagcatga 1740

Scoring table:	OLIGO_NUC	Count
	0	0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*

```

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result					Description	
No.	Score	Match	Query	Length	ID	
1	45	2.6	203	4	US-09-043-303-7	Sequence 7, Appli
2	42	2.4	397	3	US-09-253-691-3	Sequence 3, Appli
3	44	2.4	1776	2	US-08-531-9278-1	Sequence 1, Appli
4	42	2.4	1776	4	US-09-041-886-12	Sequence 12, Appli
5	41	2.4	1853	1	US-08-553-110-2	Sequence 2, Appli
6	41	2.4	4279	4	US-09-041-886-22	Sequence 22, Appli
7	38	2.2	1633	3	US-09-234-332-4	Sequence 4, Appli
8	38	2.2	3563	3	US-09-041-886-20	Sequence 20, Appli
9	38	2.2	3566	2	US-08-779-801-5	Sequence 5, Appli
10	38	2.2	3596	4	US-09-298-441-5	Sequence 5, Appli
11	38	2.2	3632	2	US-08-779-801-3	Sequence 3, Appli
12	38	2.2	3632	2	US-09-298-441-3	Sequence 3, Appli
13	37	2.1	228	6	5273901-10	Patent No. 5273901
14	37	2.1	228	6	5482709-9	Patent No. 5482709
15	37	2.1	286	2	US-08-332-766A-6	Sequence 6, Appli
16	37	2.1	510	2	US-08-134-009-150	Sequence 150, Appli
17	37	2.1	510	2	US-08-458-356-150	Sequence 150, Appli
18	37	2.1	510	4	US-08-460-736-150	Sequence 150, Appli
19	37	2.1	1332	2	US-08-481-814A-3	Sequence 3, Appli
20	37	2.1	1489	3	US-08-836-582-1	Sequence 1, Appli
21	37	2.1	1489	4	US-09-285-566-1	Sequence 1, Appli
22	37	2.1	2770	4	US-09-008-697A-13	Sequence 13, Appli
23	36	2.1	51	1	US-08-068-747-1	Sequence 1, Appli
24	36	2.1	69	4	US-09-043-303-13	Sequence 13, Appli
25	36	2.1	78	4	US-09-043-303-12	Sequence 12, Appli
26	36	2.1	154	1	US-08-469-802B-6	Sequence 6, Appli
27	36	2.1	154	2	US-08-267-803B-6	Sequence 6, Appli

28	36	2.1	165	4	US-09-043-303-17	Sequence 17, April
29	36	2.1	168	1	US-08-469-802B-4	Sequence 4, April
30	36	2.1	168	2	US-08-267-803B-4	Sequence 4, April
31	36	2.1	171	1	US-08-469-802B-5	Sequence 5, April
32	36	2.1	171	2	US-08-267-803B-5	Sequence 5, April
33	36	2.1	195	1	US-08-469-802B-2	Sequence 2, April
34	36	2.1	195	2	US-08-267-803B-2	Sequence 2, April
35	36	2.1	224	1	US-08-469-802B-3	Sequence 3, April
36	36	2.1	224	2	US-08-267-803B-3	Sequence 3, April
37	36	2.1	300	4	US-09-153-994-3	Sequence 3, April
38	36	2.1	335	2	US-08-531-927B-3	Sequence 1, April
39	36	2.1	477	4	US-09-135-994-1	Sequence 7, April
40	36	2.1	506	1	US-08-469-802B-7	Sequence 7, April
41	36	2.1	506	2	US-08-267-803B-7	Sequence 7, April
42	36	2.1	688	4	US-08-998-416-915	Sequence 915, April
43	36	2.1	715	4	US-08-998-416-590	Sequence 590, April
44	36	2.1	1080	2	US-08-918-727-2	Sequence 2, April
45	36	2.1	1080	3	US-09-205-680A-2	Sequence 2, April

ALIGNMENTS

[illegible]


```

FEATURE: NAME/KEY: CDS
; LOCATION: 239..3794
US-09-041-886-22

Query Match          2.4%; Score 41; DB 4; Length 4279;
Best Local Similarity 100.0%; Pred. No. 1.4e+09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 caacagcacagcagcagcagcagcagcagcagcagca 704
|||||
Db 1691 CAACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCA 1731

RESULT 7
US-09-234-332-4
; Sequence 4, Application US/09234332A
; Patent No. 6067168
GENERAL INFORMATION:
APPLICANT: Cedars-Sinai Medical Center
APPLICANT: Michel F. Levesque, M.D.
APPLICANT: Toomas Neuman, Ph.D.
TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
TITLE OF INVENTION: NEURONS; TRANSDIFFERENTIATION OF EPIDERMAL CELLS
FILE REFERENCE: P07 41494
CURRENT APPLICATION NUMBER: US/09/234,332A
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1635
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: gene
LOCATION: (0)..(0)
OTHER INFORMATION: Achaete scute homologous protein (ASH1) gene;
US-09-234-332-4

Query Match          2.2%; Score 38; DB 3; Length 1635;
Best Local Similarity 100.0%; Pred. No. 3.2e+08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 673 cagcagcagcagcagcagcagcagcagcagcg 710
|||||
Db 589 cagcagcagcagcagcagcagcagcagcagcg 626

RESULT 8
US-09-041-886-20
; Sequence 20, Application US/09041866
; Patent No. 6235872
GENERAL INFORMATION:
APPLICANT: Bredesen, Dale E.
APPLICANT: Rabinzadeh, Sharrooz
TITLE OF INVENTION: Proapoptotic Peptides, Dependence
TITLE OF INVENTION: Polypeptides and Methods of Use
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/09/041,886
3 FILING DATE:
4 CLASSIFICATION:
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Campbell, Cathryn A.
7 REGISTRATION NUMBER: 31,815
8 REFERENCE/DOCKET NUMBER: P-LJ 2626
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: (619) 535-9001
11 TELEFAX: (619) 535-8949
12 INFORMATION FOR SEQ ID NO: 20:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 3563 base pairs
15 TYPE: nucleic acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18 MOLECULE TYPE: DNA (genomic)
19 FEATURE:
20 NAME/KEY: CDS
21 LOCATION: 3..3550
22 US-09-041-886-20

```

Query Match 2.2%; Score 38; DB 4; Length 3563;
Best Local Similarity 100.0%; Pred. No. 3 2e-08;
Matches 38; Conservative 0; Mismatches 0; Gaps 0;

OY	673	cagcagcgacgacat cagcagcgacgacgacgcggccg	710
Dd	2970	CAGCAGCGACGACA CAGCAGCGACGACGACGCAGCC	3007

RESULT 9
 US-08-779-801-5
 Sequence 5, Application: US/08779801
 Patent No. 5853995
 GENERAL INFORMATION:
 APPLICANT: Lee, Cheng-Chi
 TITLE OF INVENTION: Large Scale Genotyping of
 NUMBER OF INVENTION: Diseases and a Diagnostic Test for Spinocerebellar Ataxia Type
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Benjamin Aaron Adler, Ph.D., J.D.
 STREET: 801 Candler Lane
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh
 SOFTWARE: Microsoft Word for Macintosh
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/779,801
 FILING DATE: January 7, 1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME/BENJAMIN AARON ADLER, Ph.D., J.D.
 REGISTRATION NUMBER: 35,442
 REFERENCE/DOCKET NUMBER: D5968
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (713) 777-2321
 TELEFAX: (713) 777-6008
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3596
 TYPE: nucleic acid
 STRANDEDNESS: double-stranded

```

?       TOPOLOGY: linear
?       MOLECULE TYPE:
?       DESCRIPTION: cDNA
?       HYPOTHETICAL: no
?       ANTI-SENSE: no
?       FRAGMENT TYPE:
?       ORIGINAL SOURCE:
?       ORGANISM: human
?       TISSUE TYPE: brain
?       IMMEDIATE SOURCE:
?       LIBRARY: primary human brain cDNA
?       CLONE: BI-1(V2)-GGCAG
?       POSITION IN GENOME:
?       CHROMOSOME/SEGMENT: 19p13
?       FEATURE:
US-08-779-801-5

```

Query Match	2.2%	Score 38	DB 2	length 3596
Best Local Similarity	100.0%	Pred. No.	3e-08	
Matches	38	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0

QY	673	710
Db	2970	3007

```

1 RESULT 10
2 US-09-298-441-5
3 ; Sequence 5, Application US/09298441
4 ; Patent No. 6303307
5 ; GENERAL INFORMATION:
6 ; APPLICANT: Lee, Cheng-Chi
7 ; TITLE OF INVENTION: Large Scale Genotyping of Diseases
8 ; TITLE OF INVENTION: and a Diagnostic Test for Sphocerebellar Ataxia Type 6
9 ; NUMBER OF SEQUENCES: 5
10 ; CORRESPONDENCE ADDRESS:
11 ; ADDRESSEE: McGregor & Adler, LLP
12 ; STREET: 8011 Candle Lane
13 ; CITY: Houston
14 ; STATE: Texas
15 ; COUNTRY: USA
16 ; ZIP: 77071
17 ; COMPUTER READABLE FORM:
18 ; MEDIUM TYPE: Floppy disk
19 ; COMPUTER: Apple Macintosh
20 ; OPERATING SYSTEM: Macintosh
21 ; SOFTWARE: Microsoft Word for Macintosh
22 ; CURRENT APPLICATION DATA:
23 ; APPLICATION NUMBER: US/09/298,441
24 ; FILING DATE:
25 ; CLASSIFICATION:
26 ; PRIOR APPLICATION DATA:
27 ; APPLICATION NUMBER: 09/093,273
28 ; FILING DATE: June 8, 1998
29 ; ATTORNEY/AGENT INFORMATION:
30 ; NAME: Benjamin Aaron Adler, Ph.D., J.D.
31 ; REGISTRATION NUMBER: 35,423
32 ; REFERENCE/DOCKET NUMBER: D5968D/C
33 ; TELECOMMUNICATION INFORMATION:
34 ; TELEPHONE: (713) 777-2321
35 ; TELEFAX: (713) 777-6908
36 ; INFORMATION FOR SEQ ID NO: 5:
37 ; SEQUENCE CHARACTERISTICS:
38 ; LENGTH: 3596 basepairs
39 ; TYPE: nucleic acid
40 ; STRANDEDNESS: double-stranded
41 ; TOPOLOGY: linear
42 ; MOLECULE TYPE:
43 ; DESCRIPTION: cdna
44 ; HYPOTHETICAL: no
45 ; ANTI-SENSE: no
46 ; ORIGINAL SOURCE:

```

ORGANISM: human
TISSUE TYPE: brain
IMMEDIATE SOURCE:
LIBRARY: primary human brain cDNA
CLONE: BI-1(V2)-GGCAG
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 19p13
US-09-298-441-5

Query Match 2.2%; Score 38; DB 4; Length 3596;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 673 cagcagcagcagcagcagcagcagcagcagcagc 710
Db 2970 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3007

RESULT 11
US-08-779-801-3
Sequence 3, Application US/08779801
Patent No. 5853995
GENERAL INFORMATION:
APPLICANT: Lee, Cheng-Chi
TITLE OF INVENTION: Large Scale Genotyping of
DISEASES AND A DIAGNOSTIC TEST FOR SPINOCEREBELLAR ATAXIA TYPE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benjamin Aaron Adler, Ph.D., J.D.
STREET: 8011 Candle Lane
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,801
FILING DATE: January 7, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benjamin Aaron Adler, Ph.D., J.D.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5968
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3632
TYPE: nucleic acid
STRANDEDNESS: double-stranded
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
TISSUE TYPE: brain
IMMEDIATE SOURCE:
LIBRARY: primary human brain cDNA
CLONE: BI-1-GGCAG
POSITION IN GENOME:

CHROMOSOME/SEGMENT: 19p13
US-08-779-801-3

Query Match 2.2%; Score 38; DB 2; Length 3632;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 673 cagcagcagcagcagcagcagcagcagcagcagc 710
Db 3006 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3043

RESULT 12
US-09-298-441-3
Sequence 3, Application US/09298441
Patent No. 6303307
GENERAL INFORMATION:
APPLICANT: Lee, Cheng-Chi
TITLE OF INVENTION: Large Scale Genotyping of Diseases
AND A DIAGNOSTIC TEST FOR SPINOCEREBELLAR ATAXIA TYPE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: McGregor & Adler, LLP
STREET: 8011 Candle Lane
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/298,441
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/093,273
FILING DATE: June 8, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Benjamin Aaron Adler, Ph.D., J.D.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5968D/C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3632 basepairs
TYPE: nucleic acid
STRANDEDNESS: double-stranded
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
TISSUE TYPE: brain
IMMEDIATE SOURCE:
LIBRARY: primary human brain cDNA
CLONE: BI-1-GGCAG
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 19p13
US-09-298-441-3

Query Match 2.2%; Score 38; DB 4; Length 3632;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 673 cagcagcagcagcagcagcagcagcagcagcagc 710
|||||
Db 3006 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3043

RESULT 13
5273901-10
; Patent No. 5273901
; APPLICANT: JACOBSON, JAMES W.; STRAUSSBERG, ROBERT L.; WILSON,
; SUSAN D.; POPE, SHARON H.; STRAUSSBERG, SUSAN L.; RUFF, MICHAEL D.;
; AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.
; TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS
; SPOROZOITE 21.5 KB ANTIGEN, AC-6B
; NUMBER OF SEQUENCES 11
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/581,693
; FILING DATE: 12-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 215,162
; FILING DATE: 05-JUL-1988
; APPLICATION NUMBER: 746,520
; FILING DATE: 19-JUN-1985
; APPLICATION NUMBER: 627,811
; FILING DATE: 05-JUL-1984
; SEQ ID NO:10:
; 5273901-10
; LENGTH:228

Query Match 2.1%; Score 37; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 9.2e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 673 cagcagcagcagcagcagcagcagcagcagcagc 709
|||||
Db 181 cagcagcagcagcagcagcagcagcagcagcagc 217

RESULT 14
5482709-9
; Patent No. 5482709
; APPLICANT: JACOBSON, JAMES W.; STRAUSSBERG, ROBERT L.; WILSON,
; SUSAN D.; POPE, SHARON H.; STRAUSSBERG, SUSAN L.; RUFF, MICHAEL D.;
; AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.
; TITLE OF INVENTION: EIMERIA ANTIGENIC COMPOSITION WHICH
; ELICITS ANTIBODIES AGAINST AVIAN COCCIDIOSIS
; NUMBER OF SEQUENCES 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,432
; FILING DATE: 08-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 581,693
; FILING DATE: 12-SEP-1990
; APPLICATION NUMBER: 215,162
; FILING DATE: 05-JUL-1989
; APPLICATION NUMBER: 746,520
; FILING DATE: 19-JUN-1985
; APPLICATION NUMBER: 627,811
; FILING DATE: 05-JUL-1984
; SEQ ID NO:9:
; 5482709-9
; LENGTH: 228

Query Match 2.1%; Score 37; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 9.2e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 673 cagcagcagcagcagcagcagcagcagcagcagc 709
|||||
Db 181 cagcagcagcagcagcagcagcagcagcagcagc 217

RESULT 15
US-08-332-766A-6
; Sequence 6, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-332-766A-6

Query Match 2.1%; Score 37; DB 2; Length 286;
Best Local Similarity 100.0%; Pred. No. 9.2e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 673 cagcagcagcagcagcagcagcagcagcagcagc 709
|||||
Db 187 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 223

Search completed: April 30, 2002, 20:40:57
Job time: 10655 sec

Wed May 1 16:36:05 2002

us-09-668-119-1.oli.rml

Page 7

1
2
3
4
5

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2002, 16:13:27 ; Search time 1517.62 Seconds
(without alignments)
12320.380 Million cell updates/sec

Title: US-09-668-119-1

Perfect score: 1740

Sequence: 1 atgaggaagctgtgtgtgc.....cacatcttcacgatga 1740

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372893281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estipl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: qb_estl:*
11: qb_estl2:*
12: qb_hlc:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	722.8	41.5	776	11	BG750043
2	697.6	40.1	730	11	BG766214
3	662	38.0	2752	12	BC005027
4	655	37.6	886	11	BG281923
5	650.8	37.4	757	10	AL046886
6	615.2	35.4	620	10	BE384305
7	572.8	32.9	939	11	BG336790
8	536.6	30.8	723	10	BE279307
9	530	30.5	1019	11	BG335275
10	527	30.3	914	11	BG468341
11	526.8	30.3	656	10	BE390284
12	524.4	30.1	526	10	BE019568

Result No.	Score	Match	Query Length	DB ID	Description
13	512.8	29.5	756	11	BG746380
14	480.4	27.6	482	10	AA664125
15	469.4	27.0	524	10	BE410770
16	469.2	27.0	733	10	BE391135
17	459.8	26.4	498	10	AW603679
18	453	26.0	942	11	BG338685
19	443.2	25.5	956	11	BG122694
20	428.8	24.6	1157	11	BG177319
21	401.6	23.1	1864	12	BC003221
22	398.2	22.9	864	11	BG386117
23	396.2	22.8	795	11	BG915202
24	386	22.2	401	10	BE163519
25	383.2	22.0	443	11	BI024741
26	377	21.7	594	10	BE284329
27	375	21.6	557	11	BG421332
28	358.2	20.6	486	10	AW747905
29	348.2	20.0	568	11	BI089323
30	343.8	19.8	690	10	BE251464
31	325.8	18.7	543	10	AW227792
32	323.8	18.6	386	11	BG982761
33	320.4	18.4	386	11	BG982754
34	310.2	17.8	416	11	BI015868
35	310	17.8	514	10	AI787536
36	307.8	17.7	487	11	BF600990
37	303	17.4	313	11	N49344
38	303	17.4	367	10	AA982166
39	301.6	17.3	335	11	BI011820
40	297.6	17.1	382	11	BG962938
41	295.4	17.0	1107	11	BF143909
42	283.4	16.3	424	10	AW501171
43	278.4	16.0	428	11	BE889171
44	278.2	16.0	435	11	BF922390
45	273	15.7	810	11	BG686794

ALIGNMENTS

RESULT 1
LOCUS BG750043 776 bp mRNA
DEFINITION 602708813F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4845411 5', mRNA sequence.
ACCESSION BG750043
VERSION BG750043.1 GI:14060696
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 776)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNML Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L16M1684 row: j column: 04
High quality sequence stop: 743.
Location/Qualifiers
1..776
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:4845411"
/clone_lib="NIH_MGC_43"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"

Db	362	CGTCCCGACGCCGGCGAGCCCGAGCTCACAGCCCAACTCAAGTCAAGCTGTGGCCTTG	421
QY	1223	ccccatctccagtagcttctctgcacagcccctcaaccgacgcctctccagagcccagttga	1282
Db	422	CCCCATCTCCACAGTAGGTTCTGTGCCACGCCCTCACCGCAGCCCTCCCAAGGCCAATGA	481
QY	1283	cgagcgagaccccaacagaacttcagttgcctctcaacttgagaccttaaacacacttga	1342
Db	482	CGGCGCGGAGCCCAACAGACTTCAGTGTCCCTTACTGTGACCTTTAAACACACTGTGA	541
QY	1343	accccaactctgtcaatgagcccactgctgcctcagccagctgtagaagcaagtaacctg	1402
Db	542	ACCCCACTCTGTATGATGAGCCCAACTGGCTCCACCCAGGCTGAGAGACACAGTAACTGG	601
QY	1403	acaagctggaagcacctgttgaagtaactgaagcccctggcgcgatataaagaagtgc	1462
Db	602	ACAAGCTGAAACACTGTCTGAAGTACATCGAGCCCTCGCCGGCATATCAACAAGATCG	661
QY	1463	acaaaga-cgaagaacaga-aaaaagagacctgtagtgaagaatgaagacctcttgcacattc	1520
Db	662	ACAGAAGACCGAAGACAGACCAAAAGGACCTGATGAAGTGAAGAGGCTTCTGGACATTCT	721
QY	1521	gacagacc 1528	
Db	722	GACAGAAC 729	

```

VERSION      BG281923.1  31:13030849
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
              Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
REFERENCE    1 (bases 1 to 886)
              NIH-MGC http://mgi.nci.nih.gov/
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
AUTHORS      Contact: Robert Strausberg, Ph.D.
              Email: cgapds@mail.nih.gov
              Tissue Procurement: ATCC/DC/DTP
              CDNA Library Preparation: Ling Hong/Rubin Laboratory
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
              Plate: LNCM230 row: 9 column: 02
              High quality sequence stop: 876.
FEATURES
  source
    1..886
    Location/Qualifiers
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:454533"
      /clone_1b="NIH_MGC_20"
      /issue_type="melanotic melanoma"
      /lab_host="DH10B (phage-resistant)"
      /note="Organ: skin; Vector: pOT7; Site_1: XhoI; Site_2:
      EcoRI; CDNA made by oligo-dT priming. Directionally
      cloned into EcoRI/XhoI sites using the following 5'
      adaptor: GGCACAGAG(G). Size selected >500bp for average
      insert size 1.8kb. Library constructed by Ling Hong in
      the laboratory of Gerald M. Rubin (University of
      California, Berkeley) using ZAP-cDNA synthesis kit
      (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT   234 a      299 c      231 g      122 t
ORIGIN
```

```

Query Match      37.6%; Score 655; DB 11; Length 886;
Best Local Similarity 93.7%; Pred. No. 1.7e-98;
Matches 749; Conservative 0; Mismatches 40; Indels 10; Gaps 6;

QY 291 gttctctcaggagcgcgcctctcctggaactcgggagatggccctcaacatgagctgt 350
    |||||||
Db 2 gttctctcaggagcgcgcctctcctggaactcgggagatggccctcaacatgagctgt 61
    |||||||

QY 351 cgtgtctacggcaagtccacacacacagctgcagctcagcaagtgagctgacagca 410
    |||||||
Db 62 cgtgtctacggcaagtccacacacacagctgcagctcagcaagtgagctgacagca 121
    |||||||

QY 411 gtcagcaacagcagcagttccacagcagcagcagcagcagcagcagcagcagca 470
    |||||||
Db 122 GCAGCAACACAGCAGCAGTTCCAGCAGCAGCAGCAGCGCGCTACAGCAGCAGCAGCA 181
    |||||||

QY 471 gtcagcaacagcagcagttccacagcagcagcagcagcagcagcagcagcagc 530
    |||||||
Db 182 GCAGCAACACAGCAGCAGTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAT 241
    |||||||

QY 531 agtcagcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 590
    |||||||
Db 242 AGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAT 301
    |||||||

QY 591 gtcagcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 650
    |||||||
Db 302 GCATCATCAAAATCAAGCAACAGATACAGCAGCAGCAACAGCAGCTCAGCAGATACACA 361
    |||||||

QY 651 gtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 706
    |||||||
Db 362 GCTGCAAGCTTCACACACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 421
    |||||||

QY 707 aggccttgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 765
    |||||||
```

```

Db 422 ggctttggcagggccagccagccacacaaatttcagccagccagcagcagcagcagc 481
    |||||||
QY 766 cctccagagcgtctgccccagcagcagcagcagcagcagcagcagcagcagcagcagc 825
    |||||||
Db 482 TCCTCCAGAGCTCTGCCAGCAGCTGCAGAGATGATACACACAGCAGCAGCAGCAGCAG 541
    |||||||
QY 826 ccaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 884
    |||||||
Db 542 -CACCAAGCCCCAGCAGCAGCTCAGTGTCTGAAAGCAAGCATGCAACAATCCGCCGAAA 600
    |||||||
QY 885 gtcgagagccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 944
    |||||||
Db 601 GTCGAGAGCCAGCAGCTTGGTGTCTACAGGCGCAAGCTCTCCCTGACAAATGTTGATAC 660
    |||||||
QY 945 ccaaccacactgaattgtccagagctcc-gatggtgtgcaagcag--ccccagtgca 1001
    |||||||
Db 661 CCAAGCAGCAATGAATTTGTCCGAGCTCTGATGTGTGTGTCAGCAGGCCCCAGTGCCA 720
    |||||||
QY 1002 gccccagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1061
    |||||||
Db 721 GCACAGGTGACAGCAGCAGGAAGACAGCAGTACAGACCTGCTCAGCTGCCAGATGCTGC 780
    |||||||
QY 1062 tcccgagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
    |||||||
Db 781 TCCTCGAATCCAGATGATC 799
    |||||||

RESULT 5
AL046886
LOCUS
DEFINITION
  DKFZP586E2117_r1 586 (synonym: hute1) Homo sapiens CDNA clone
  DKFZP586E2117 5', mRNA sequence.
ACCESSION
  AL046886
VERSION
  AL046886.1 GI:5434945
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
  Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
REFERENCE
  1 (bases 1 to 757)
AUTHORS
  Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
  EST (Koehler, 1999)
JOURNAL
  Unpublished (1999)
COMMENT
  Contact: Koehler K
  MFS
  Am Klopferspitz 18a D-82152 Martinsried, Germany
  This is the 5' sequence of the clone insert
  This is the 5' sequence of the clone insert
  Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
  Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
  sequenced by BMFZ (Biomedical Research Center at the Charite,
  Berlin/Germany) within the cdna sequencing consortium of the German
  Genome Project.
  No sl sequence available.
  This clone (DKFZP586E2117) is available at the RZPD in Berlin.
  Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
  Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
  source
    1..757
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="DKFZ586E2117"
    /clone_1b="586 (synonym: hute1)"
    /issue_type="uterus"
    /dev_stage="adult"
    /lab_host="DH10B"
    /note="Vector: pSport1; site_1: NotI; site_2: SalI/MluI"
BASE COUNT   199 a      233 c      208 g      116 t
ORIGIN
```

```

Query Match      37.4%; Score 650.8; DB 10; Length 757;
Best Local Similarity 98.0%; Pred. No. 8.2e-98;
```



```

ACCESSION      BG468341
VERSION        BG468341.1 (1:13400611)
KEYWORDS
SOURCE
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 914)
                NIH-MGC http://mgi.nci.nih.gov/
                National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cga@r-remail.nih.gov
                Tissue Procurement: ATCC
                CDNA Library Preparation: Ling Hong/Rubin Laboratory
                DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LICM1415 row: n column: 12
                High quality sequence stop: 812.
                Location/Qualifiers
                1. 914
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:464299"
                /clone_lib="NIH-MGC_15"
                /tissue_type="adenoecarcinoma cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: colon; Vector: pORF7; Site:1: XhoI; Site:2:
                EcoRI; CDNA made by oligo-dt priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGCACGAG(G). Size-selected >500bp for average
                insert size 1.8kb. Library constructed by Ling Hong in
                the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies)."
                (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT    225 a 305 c 229 g 155 t
ORIGIN
Query Match   30.3%; Score 527; DB 11; Length 914;
Best Local Similarity 99.8%; Pred. No. 2e-77;
Matches 538; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1203 caacgtcagctctgacctgcccacatctccacagtaactctctccacagccctccacagca 1262
      |||
DB 3 CAACGTGAGCTCTG3CCCTGCCCCATCTCCAGTAGCTTCTGCCAGCCCTCCACCGCA 62

QY 1263 gccctccacagagccagtgacgagcgagccacacagaaactcagtgctccctcacctg 1322
      |||
DB 63 GCCCTCCAGAGCCGAGTGAGCGGGCGGAGCCCAAGAACTTCAGTGTCCCTCACCTG 122

QY 1323 accttaaacacacgttgtaaccacagctctgcatgagccagctggtctccacagca 1382
      |||
DB 123 ACCTTTAAACACACGTGTGAACCCAGCTCTGTATGAGCCGACGTGGCTCCACGACG 182

QY 1383 ttagagagcagcagtgactgtaaacagctgagcagctgtgaataatcagagccctgag 1442
      |||
DB 183 TGAAGAGCAGCAGTACCTTGACAGAGCCCTCGAAGCGGTGTCCCTGAAGACCTTGCAAAA 242

QY 1443 ccgcatgatcaacaagatcgacaagaagaagaagaagaagaagaagaagaagaagaaga 1502
      |||
DB 243 CCGCATGATCAACAAGATCGACAAGAAAGCAAGACAGAAAAAGAGACTGAGTAGAGAA 302

QY 1503 gaggctcttgagaca:ctctgacagaccctcgaagcggtgtccctcgaagccttgaaaa 1562
      |||
DB 303 GAGCCTTCTGAGACAATCTGACAGAGCCCTCGAAGCGGTGTCCCTGAAGACCTTGCAAAA 362

QY 1563 gtgtgagatcgccc:agagaagaactcaagaatgacat-gcggtgcccactcccccacgccc 1621
      |||
DB 363 GTGTGAGATCGCCCTGGAAGAAATCAAGATGACATGGCGGTGCCACACTCCACACGCC 422

```

```

QY 1622 cccggtgccaccagcaaacacagcagtaactatgccaagccgctctgagatgcgctcgc 1681
      |||
DB 423 CCCGGTCCACCCAGACCAAGACAGCAGTACTATGCGACGCGCTCTGATGCGCTGCGC 482

QY 1682 caacatcgctcagcagctgctcacaacatccctgtacagcagcattgctcagcagatga 1740
      |||
DB 483 CAACATCGCTCAGCAGCTGTCTTCAACCATCTCTTACCGACATCTGCTCCAGCAGCA 541

RESULT 11
LOCUS        BE390284      656 bp      mRNA      EST      21-JUL-2000
DEFINITION  601285167F1 NIH-MGC_44 Homo sapiens cDNA clone IMAGE:3606905 5',
ACCESSION   BE390284
VERSION     BE390284
KEYWORDS    BE390284.1 GI:9335562
SOURCE      human.
ORGANISM    Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 656)
                NIH-MGC http://mgi.nci.nih.gov/
                National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cga@r-remail.nih.gov
                Tissue Procurement: ATCC
                CDNA Library Preparation: Ling Hong/Rubin Laboratory
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LICM257 row: e column: 18
                High quality sequence stop: 588.
                Location/Qualifiers
                1. 656
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:3606905"
                /clone_lib="NIH-MGC_44"
                /tissue_type="endometrium, adenoecarcinoma cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: uterus; Vector: pORF7; Site:1: XhoI; Site:2:
                EcoRI; CDNA made by oligo-dt priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGCACGAG(G). Library constructed by Ling Hong
                in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies)."
                (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT    181 a 239 c 157 g 79 t
ORIGIN
Query Match   30.3%; Score 526.8; DB 10; Length 656;
Best Local Similarity 94.0%; Pred. No. 2.1e-77;
Matches 613; Conservative 0; Mismatches 32; Indels 7; Gaps 6;

QY 538 cagcagcagcagctcacaagcagcagcagcagcagcagcagcagcagcagcagcagcag 597
      |||
DB 1 CAGCGGCGACGAGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 60

QY 598 caaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 657
      |||
DB 61 CAAAATCAGCAACATATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120

QY 658 ctccaacaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 717
      |||
DB 121 CTCCAACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 180

QY 718 gccacgacacacattcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 777

```


Fri Mar 1 09:12:00 2002

us-09-668-119-1.rst

Page 12